

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-267937

(43)Date of publication of application : 29.09.2000

(51)Int.Cl.

G06F 12/10  
G06F 15/167

(21)Application number : 11-073068

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 18.03.1999

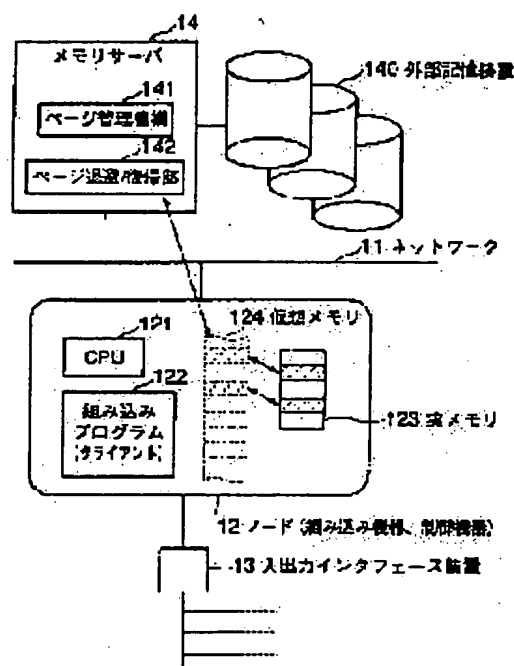
(72)Inventor : MIURA MASAHIRO

## (54) VIRTUAL STORAGE CONTROL SYSTEM

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To virtually supply a lot of memories to a built-in program to be operated in an electronic equipment from the outside through a network.

**SOLUTION:** In this virtual storage control system, a node 12 incorporated with a CPU 121 for executing a built-in program 122 using a virtual storage is connected to a memory server 14 provided with an external storage device for storing contents of memory pages in the virtual storage, a page management mechanism 141 and a page save/return part 142 through a network. When a request for saving or returning a memory page in the virtual storage is generated from the built-in program (client) 122, the memory server 14 is informed of the request through the network 11 and the save or return processing of the requested memory page is executed by the page save/return part 142 in the server 14 by using the external storage device 140.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Japanese Patent Laid-open Publication No. 2000-267937 A

Publication date: September 29, 2000

Applicant:: TOSHIBA CORP

Title: Virtual storage management system

5

(57) [Abstract]

[Object] To make it possible to virtually provide a large quantity of memory externally via a network, with respect to a built-in program operating in electronic equipment.

[Means] The configuration is such that a node 12 having a built-in CPU 121 that  
10 executes a built-in program 122 using a virtual storage, and a memory server 14 having an external storage unit 140 that stores the entity of a memory page of the virtual storage, a page management mechanism 141, and a page saving/return section 142 are connected to each other by a network 11, and when a saving or return request of the memory page in the virtual storage is made from the built-in program  
15 (client) 122, the request is notified to the memory server 14 via the network 11, and the saving or return processing of the requested memory page is performed by the page saving/return section 142, by using the external storage unit 140.

[What is claimed is]

20 [Claim 1] A virtual storage control system comprising:

first equipment having a CPU built therein, which executes a built-in program, using a virtual storage;

second equipment including a large capacity memory in which the entity of the memory page in the virtual storage is stored, and a page management unit that  
25 manages the respective memory pages; and

a network to which the first and the second equipment are connected,  
wherein the first equipment comprises a page saving/return request unit that  
transmits a saving or return request of a memory page to the second equipment via  
the network, when the saving or return request of the memory page in the virtual  
5 storage is made from the built-in program operating in the equipment, and the second  
equipment comprises a page saving/return unit that performs the saving or return  
processing of the requested memory page by using the memory, in response to the  
request from the page saving/return request unit in the first equipment.

[Claim 2] The virtual storage control system according to claim 1, wherein the  
10 first equipment is constructed so that a plurality of processes can be operated in  
parallel, and

the page management unit in the second equipment manages the respective  
memory pages, together with identification information of the process using the page.

[Claim 3] The virtual storage control system according to claim 2, wherein a  
15 plurality of the first equipment is connected to the network, and  
the page management unit in the second equipment manages the respective  
memory pages, together with identification information of the first equipment using the  
page.

[Claim 4] The virtual storage control system according to claim 1 or 2, wherein  
20 a plurality of the second equipment is connected to the network, and

the first equipment further comprises a page saving destination/return source  
determination unit that notifies all the second equipment on the network of a request for  
a saving destination or a return source of the memory page by broadcast, in order to  
determine the saving destination or the return source of the memory page, and  
25 determines the saving destination or the return source of the memory page, based on

the response with respect to the notification.

[Claim 5]           The virtual storage control system according to claim 1 or 2, wherein a plurality of the second equipment is connected to the network, and third equipment that determines the equipment to be a saving destination or a return source of the

5   memory page from the second equipment is also connected thereto, and

                  the first equipment further comprises a page saving destination/return source determination unit that inquires of the third equipment via the network, when it is necessary to determine the saving destination or the return source of the memory page, and determines the saving destination or the return source of the memory page

10   based on the response with respect to the enquiry.

[0005]

[Problems to be solved by the Invention]

                  Recently, the function required for the built-in program is upgraded.

15   Accompanying the up-grading, the program becomes more complicated, and hence, high maintainability of the program and easiness in reusing the program parts are required.

[0006]   Highly developed programs require a large quantity of memory, and a large quantity of memory and a highly developed memory management mechanism (virtual storage and the like) are also required in order to use a language, which can be described by a program having high maintainability and reusability (such as Java developed by Sun Microsystems in USA).

20                   [0007]   However, conventional electronic equipment (built-in equipment) such as management equipment in which the built-in program operates has a strict restriction in  
25   the quantity of the memory that can be used, due to the production cost and the like.

Therefore, in this type of conventional electronic equipment, it is difficult to apply a highly developed built-in program.

[0008] The present invention has been achieved in order to solve the above problems. It is an object of the present invention to provide a virtual storage system  
5 that can virtually supply a large quantity of memory to the built-in program operating in the electronic equipment externally via a network, thereby enabling application of the highly developed built-in program.

[0017]

#### 10 [Embodiments of the Invention]

The embodiments of the present invention will be explained with reference to the accompanying drawings.

[0018]

[First Embodiment]

15 Fig. 1 is a block diagram illustrating the configuration of a virtual storage control system according to the first embodiment of the present invention. In this figure, electronic equipment (built-in equipment) having a built-in CPU 121, for example, management equipment (hereinafter, node) 12 is connected to a high-speed network  
11 such as Ethernet, Asynchronous Transfer Mode (ATM), or Fibre Channel. Various  
20 objects to be managed (not shown) are connected to the node 12 via an input and output interface device 13.

[0019] In the node, the CPU 121 executes a built-in program 122 that uses a virtual storage (and comprises a built-in OS and a program operating under management of the OS). The node 12 has a built-in small-capacity physical memory (actual memory)  
25 123. In the node 12, a virtual memory 124 having a large capacity exists equivalently,

which can be seen from the built-in program 122. This virtual memory 124 is provided by a memory server 14 described below. The data of a memory page (actual page data) in the virtual memory 124, to be processed in the node 12, is arranged in an optional area on the actual memory 123, so that it can be processed by the CPU 121 in the node 12.

[0020] To the network 11 is also connected a memory server (memory server equipment) 14 realized by information processing equipment such as a personal computer or a workstation. The memory server 14 is for providing a virtual memory 124 (upper page) to the built-in program 122 on the node 12, and includes a large-capacity external storage unit 140 that stores the entity of the page on the virtual memory 124. The external storage unit 140 can be realized by a disk array using hard disks, as shown in the example of Fig. 1, as well as by a semiconductor disk, a DVD-RAM and the like. The memory server 14 has a page management mechanism 141 that manages memory pages stored (saved) in the external storage unit 140, and memory pages provided to (return in) the built-in program 122 on the node 12, and a page saving/return section 142 that executes saving or return processing of the memory page requested from the built-in program 122 in the node 12, by using the external storage unit 140.

[0021] The operation in the system having the configuration shown in Fig. 1 will be explained next. It is assumed here that saving and return of the page (memory page in the virtual storage) on the virtual memory 124 become necessary, by the operation of the built-in program 122 on the node 12. In the case of the general information processing equipment such as a personal computer applying the virtual storage method, since it has a large-capacity memory such as a hard disk or a semiconductor disk as a part of the equipment, saving and return of the memory page is performed

with respect to the memory. However, since the node 12 as the built-in equipment in this embodiment does not have a large-capacity memory, it cannot perform saving and return of the memory page, as performed by the general information processing equipment.

5 [0022] Therefore in this embodiment, the built-in program 122 operating in the node 12 performs saving and return of the memory page by requesting it to the external memory server 14. In other words, the built-in program 122 has a function of requesting saving and return of the memory page. Therefore, hereinafter, the built-in program 122 in the node 12 is referred to as a client 122.

10 [0023] When a memory page is for example saved, the client 122 side transmits the memory page information having the structure as shown in Fig. 2, that is, the memory page information including the contents 21 of the page to be saved (actual page data) 21 and page management information 22 including an address (virtual address) VA of the page on the virtual memory 124, from the node 12 to the memory server 14 via the  
15 network 11. The memory server 14 side stores the received memory page information in the external storage unit 140 by the page saving/return section 142. Alternatively, only the actual page data 21 of the memory page information may be stored in the external storage unit 140, and the page management information 22 may be stored and managed by the page control mechanism 141.

20 [0024] When the memory page is returned, the client 122 side transmits a virtual address VA of the page to be returned from the node 12 to the memory server 14 via the network 11, to request transmission of the contents of the page.

[0025] The memory server 14 specifies the memory page requested from the client 11 by the page management mechanism 141, based on the virtual address VA  
25 transmitted from the client 122. Then, the memory server 14 fetches the information

of the requested memory page from the external storage unit 140 by the page saving/return section 142, and transmits the information to the requested client 122 via the network 11. The client 122 arranges the information of the requested memory page transmitted from the memory server 14 in an optional area on the actual memory 123, so that the CPU 121 can process the information in the node 12. Thereby, the return processing of the memory page by the cooperation operation with the memory server 14 in the node 12 is finished.

[0026] When the client 122 has a parallel execution function of a plurality of processes, the respective processes have the respective virtual memories (virtual spaces) virtual memory 124 on the node 12. In this case, the information of the memory page transferred between the memory server 14 and the client 122 may include page management information 32 in which information for identifying the process (process ID) is added, as shown in Fig. 3, instead of the page management information 22 in Fig. 2.

[0027]

[Second Embodiment]

In the first embodiment, an example of a system in which one memory server and one client exist is shown. However, the present invention is also applicable to a system in which a plurality of memory servers and a plurality of clients exist.

Therefore, the second embodiment, in which the present invention is applied to a virtual storage control system including a plurality of memory servers and a plurality of clients, will be explained with reference to the accompanying drawings.

[0028] Fig. 4 is a block diagram illustrating the configuration of a virtual storage control system according to the second embodiment of the present invention. In this figure, a plurality of built-in equipment, for example, two built-in equipment (nodes) 42-1



and 42-2, and a plurality of memory servers, for example, two memory servers 44-1 and 44-2 are respectively connected to the network 41 (corresponding to the network 11 in Fig. 1).

[0029] The nodes 42-1 and 42-2 respectively include a CPU 421-1, 421-2, a built-in  
5 program (client) 422-1, 422-2, and a small-capacity actual memory 423-1, 423-2 built therein. In the nodes 42-1 and 42-2, a virtual memory 424-1, 424-2, which can be seen from the client 422-1, 422-2, respectively exists equivalently for the number of processes that can be parallel executed by the clients 422-1 and 422-2. The virtual memory 424-1, 424-2 is provided respectively by a memory server 44-1, 44-2.

10 [0030] In the system shown in Fig. 4, since a plurality of clients (here, the client 422-1 in the node 42-1, and the client 422-2 in the node 42-2) exists, the memory page information transferred between the memory server 44-1, 44-2 and the client 422-1, 422-2 also includes information for identifying the node (client) (node ID). In the system of Fig. 4, since the respective processes have the respective virtual memories  
15 (virtual spaces) 424-1 and 424-2 on the nodes 42-1 and 42-2, information for identifying the process (process ID) is also required as in Fig. 3. In this embodiment, therefore, page management information 52 in which the node ID is added, as shown in Fig. 5, is used, instead of the page management information 32 shown in Fig. 3.

[0031] In the system in Fig. 4, since a plurality of memory servers exist (here, two  
20 memory servers, 44-1 and 44-2), the system requires a mechanism for determining which server becomes the memory page saving destination from a certain client. Two methods can be applied for the determination of the memory page saving destination, that is, a first saving destination determination method in which the memory server to be the saving destination is determined in advance one for each  
25 client, and a second saving destination determination method in which the memory

server to be the saving destination is dynamically determined every time the memory page is saved.

[0032] Here, the procedure of the first saving destination determination method will be explained, with reference to the sequence chart in Fig. 6, using an example in which  
5 a memory page saving destination from the client 422-1 in the node 42-1 is determined.

[0033] The client 422-1 notifies all memory servers on a network 41 simultaneously of a request for a memory server that becomes the saving destination of the memory page (step S1). This notification can be easily realized, by using the mechanism of  
10 broadcast, if the network 41 applies a protocol referred to as an Internet Protocol (IP), for example, in Ethernet.

[0034] The respective memory servers 44-1 and 44-2 give a response corresponding to the respective situations (saving destination server notification or non-saving destination server notification) to the client 422-1. Here, the memory  
15 server 44-2 returns a saving destination server notification notifying that the memory server 44-2 becomes the saving destination server for the memory page from the client 422-2 (step S2), and the memory server 44-1 returns a non-saving destination server notification notifying that the memory server 44-1 does not become the saving  
destination server for the memory page from the client 422-2 (step S3).

[0035] Upon reception of these notifications, the client 422-1 selects a memory  
20 server as the saving destination of the memory page. Here, the memory server 44-2 is selected as the saving destination. Thereafter, the client 422-1 transmits the memory page information having the structure as shown in Fig. 5 from the node 42-1 to the memory server 44-2, to request saving of the corresponding memory page,  
25 every time it becomes necessary to save the memory page.

[0036] In the above explanation, not only the memory server that becomes the saving destination of the memory page, but also the memory server that does not become the saving destination gives a response, but the configuration may be such that the memory server that does not become the saving destination does not give a response.

[0037] The second saving destination determination method will be explained. In the second saving destination determination method, the request for the saving destination memory server by broadcast from the client 422-1 to the respective servers, which is applied in the first saving destination determination method, is transmitted every time the saving of the memory page becomes necessary, not in advance. Therefore, in the second saving destination determination method, the memory server that becomes the saving destination is determined every time the memory page is to be saved.

[0038] As shown in Fig. 7, it is also possible to have a configuration such that an arbitrator 70, which is a special server having a function of determining a pair of a client that becomes a saving source (or a return destination) of the memory page and a memory server that becomes a saving destination (or a return source), is added to the system in Fig. 4, so that the respective clients 422-1 and 422-2 inquire of the arbitrator 70 about a memory server that becomes the saving destination (or the return source) of the memory page.

[0039] The same method as in the first saving destination determination method, that is, the method in which the client requests the arbitrator 70 to determine the memory server to be the saving destination in advance, and the same method as the second saving destination determination method, that is, the method in which every time saving of the memory page becomes necessary, the client requests the arbitrator 70 to

dynamically determine the memory server to be the saving destination, can be applied to the system of Fig. 6.

[0040] Return of the memory page in the system in Fig. 4 or Fig. 6 will be explained, using an example when the client 422-2 returns the memory page.

5 [0041] At first, in the system in Fig. 4 or Fig. 6, since a plurality of memory servers exists, when the client 422-2 returns the memory page, it is necessary to determine from which server the memory page is to be returned, that is, determine the return source memory server, as in the instance when the memory server to be the saving destination is determined. Either one of methods can be applied for the determination  
10 of the return source memory server of the memory page, that is, a first return source determination method in which it is set such that return is always performed from the saving destination server, and a second return source determination method in which every time the return of the memory page becomes necessary, the memory server to be the return source is dynamically determined.

15 [0042] The first return source determination method has the simplest processing, and the client 422-1 needs only to memorize the memory server that becomes the saving destination. When this method is applied, however, there is a problem in that a large load is applied on a specific memory server, according to the system situation.

[0043] The second return source determination method is to solve the above  
20 problem in the first return source determination method, and it is based on the assumption that the memory page moves between memory servers according to the load status and the like, that is, the saving destination memory server does not always agree with the return source memory server, in order to realize load balancing between the memory servers.

25 [0044] When the second return source determination method is applied to the

system in Fig. 4, every time the return of the memory page becomes necessary, a request for the return source memory server is transmitted by broadcast to all memory servers (44-1, 44-2) on the network 41, by the same method as the second saving destination determination method, to determine the return source memory server  
5 based on the response thereto.

[0045] When the second return source determination method is applied to the system in Fig. 6, every time the return of the memory page becomes necessary, the client may inquire of the arbitrator 70 about the return source memory server, and determine the return source memory server based on the response thereto.

10 [0046]

[Effects of the Invention]

According to the present invention, in the built-in equipment, a large quantity of memory can be used by the virtual storage, without providing an external storage unit such as a hard disk.

15 [Brief Description of the Drawings]

[Fig. 1] A block diagram illustrating the configuration of a virtual storage control system according to the first embodiment of the present invention.

[Fig. 2] A diagram illustrating an example of data structure of memory page information applied in the embodiment.

20 [Fig. 3] A block diagram illustrating an example of data structure of memory page information, when a client (built-in program) parallel-executes a plurality of processes.

[Fig. 4] A block diagram illustrating the configuration of a virtual storage control system according to the second embodiment of the present invention.

[Fig. 5] A diagram illustrating an example of data structure of memory page  
25 information applied in the embodiment.

[Fig. 6] A sequence chart for explaining the procedure of a first saving destination determination method applied in the embodiment, using an example in which a memory page saving destination from a client 422-1 in a node 42-1 is determined.

[Fig. 7] A block diagram illustrating a modification example of the configuration shown in Fig. 4.

[Description of Signs]

11, 41 ... Network

12, 42-1, 42-2 ... Node (built-in equipment, first equipment)

14, 44-1, 44-2 ... Memory server (second equipment)

10 22, 32, 52 ... Page management information

70 ... Arbitrator (third equipment)

121, 421-1, 421-2 ... CPU

122, 422-2, 422-2 ... Built-in program (client, page saving/return request unit, page saving destination/return source determination unit)

15 123, 423-1, 423-2 ... Actual memory

124, 424-1, 424-2 ... Virtual memory

141 ... Page management mechanism

142 ... Page saving/return section

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-267937

(P2000-267937A)

(43) 公開日 平成12年9月29日 (2000.9.29)

(51) Int.Cl.<sup>7</sup>

識別記号

F I

テーマコード(参考)

G 0 6 F 12/10

G 0 6 F 12/10

F 5 B 0 0 5

15/167

15/167

Z 5 B 0 4 5

審査請求 未請求 請求項の数 5 O L (全 7 頁)

(21) 出願番号 特願平11-73068

(22) 出願日 平成11年3月18日 (1999.3.18)

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 三浦 雅弘

東京都府中市東芝町1番地 株式会社東芝

府中工場内

(74) 代理人 100058479

弁理士 鈴江 武彦 (外6名)

Fターム(参考) 5B005 JJ21 KK02 MM31

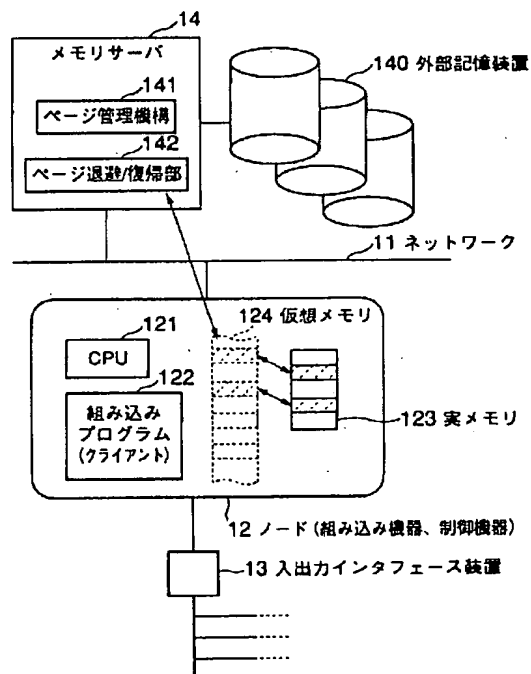
5B045 BB47 DD11

(54) 【発明の名称】 仮想記憶制御システム

(57) 【要約】

【課題】 電子機器内で動作する組み込みプログラムに対し、外部からネットワークを介して仮想的に大量のメモリが供給可能なようにする。

【解決手段】 仮想記憶を用いた組み込みプログラム122を実行するCPU121を内蔵したノード12と、仮想記憶のメモリページの実体が記憶される外部記憶装置140、ページ管理機構141及びページ退避/復帰部142を備えたメモリサーバ14とをネットワーク11により相互接続し、組み込みプログラム(クライアント)122から仮想記憶のメモリページの退避または復帰の要求が発生した場合に、その要求をネットワーク11経由でメモリサーバ14に通知して、要求されたメモリページの退避または復帰処理を当該メモリサーバ14のページ退避/復帰部142により外部記憶装置140を用いて行う構成とする。



## 【特許請求の範囲】

【請求項1】 仮想記憶を用いた組み込みプログラムを実行するCPUを内蔵した第1の機器と、

仮想記憶のメモリページの実体が記憶される大容量の記憶装置、及び前記各メモリページを管理するページ管理手段を備えた第2の機器と、

前記第1及び第2の機器が接続されたネットワークとを具備し、

前記第1の機器は、当該機器内で動作する前記組み込みプログラムから仮想記憶のメモリページの退避または復帰の要求が発生した場合に、そのメモリページの退避または復帰要求を前記ネットワーク経由で前記第2の機器に送出するページ退避／復帰要求手段を備えると共に、前記第2の機器は、前記第1の機器の前記ページ退避／復帰要求手段からの要求に応じて、要求されたメモリページの退避または復帰処理を前記記憶装置を用いて行うページ退避／復帰手段を備えていることを特徴とする仮想記憶制御システム。

【請求項2】 前記第1の機器は、複数のプロセスが並列動作可能なように構成されており、

前記第2の機器の前記ページ管理手段は、前記各メモリページを、当該ページを利用する前記プロセスの識別情報と共に管理することを特徴とする請求項1記載の仮想記憶制御システム。

【請求項3】 前記ネットワークには前記第1の機器が複数接続されており、

前記第2の機器の前記ページ管理手段は、前記各メモリページを、当該ページを利用する前記第1の機器の識別情報と共に管理することを特徴とする請求項1または請求項2記載の仮想記憶制御システム。

【請求項4】 前記ネットワークには前記第2の機器が複数接続されており、

前記第1の機器は、メモリページの退避先または復帰元を決定するために、メモリページの退避先または復帰元を要求していることを、一斉同報により前記ネットワーク上の全ての前記第2の機器に通知して、その通知に対する応答に基づいて、メモリページの退避先または復帰元を決定するページ退避先／復帰元決定手段を更に備えていることを特徴とする請求項1または請求項2記載の仮想記憶制御システム。

【請求項5】 前記ネットワークには前記第2の機器が複数接続されると共に、前記複数の第2の機器の中からメモリページの退避先または復帰元となる機器を決定する第3の機器が接続されており、

前記第1の機器は、メモリページの退避先または復帰元を決定する必要がある場合、前記第3の機器に前記ネットワークを介して問い合わせ、その問い合わせに対する応答に基づいて、メモリページの退避先または復帰元を決定するページ退避先／復帰元決定手段を更に備えていることを特徴とする請求項1または請求項2記載の仮

想記憶制御システム。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、組み込みプログラムが動作する電子機器を備えたシステムに係り、特に当該機器で大量に仮想的なメモリを利用可能な仮想記憶制御システムに関する。

## 【0002】

【従来の技術】 一般に、制御機器などに組み込まれたCPUが実行する組み込みプログラム、即ち組み込みOS（オペレーティングシステム）及び当該OSの管理下で動作するプログラムでは、利用できるハードウェア資源がパーソナルコンピュータ等に比べて乏しくなっている。その理由は、製造コストが極力抑えられることや、ハードウェアが動作する環境が（温度や振動などの点で）苛酷であることなどにある。

【0003】 このような条件から、組み込みプログラムが使えるメモリの量にも厳しい制限がある。したがって、パーソナルコンピュータ等で一般的なOSが快適に動作できるだけの容量のメモリを、組み込みプログラムが動作する制御機器（組み込み機器）で確保できる場合は稀である。また、機器にはハードディスク等の外部記憶装置は付いていない場合が多いので、仮想記憶も利用できない。

【0004】 このようなメモリ量の制約から、従来の組み込み機器における組み込みプログラムはアセンブラ（或いは、せいぜいC言語）で記述され、メモリ操作は絶対アドレスを直接指定して行っていた。

## 【0005】

【発明が解決しようとする課題】 近時、組み込みプログラムに求められる機能が高度化してきている。また、それに伴ってプログラムが複雑化することから、プログラムの保守性の高さやプログラム部品の再利用のしやすさなどが求められるようになっている。

【0006】 高度なプログラムにはそれなりに大量のメモリが必要であり、また保守性や再利用性の高いプログラムが記述可能な言語（米国サン・マイクロシステムズ社が開発したJavaなど）を使うためにも、やはり大量のメモリや、高度なメモリ管理機構（仮想記憶など）が必要となる。

【0007】 ところが、組み込みプログラムが動作する従来の制御機器等の電子機器（組み込み機器）には、前記したように製造コスト等の関係で使えるメモリの量に厳しい制限があった。したがって、この種の従来の電子機器では、高度な組み込みプログラムを適用することは困難であった。

【0008】 本発明は上記事情を考慮してなされたものでその目的は、電子機器内で動作する組み込みプログラムに対し、外部からネットワークを介して仮想的に大量のメモリを供給することができ、これにより高度な組み



込みプログラムの適用を可能とする仮想記憶システム。

【0009】

【課題を解決するための手段】本発明は、仮想記憶を用いた組み込みプログラムを実行するCPUを内蔵した第1の機器と、仮想記憶のメモリページの実体が記憶される大容量の記憶装置、及び前記各メモリページを管理するページ管理手段を備えた第2の機器とをネットワークにより相互接続し、上記第1の機器には、当該機器内で動作する組み込みプログラムから仮想記憶のメモリページの退避または復帰の要求が発生した場合に、そのメモリページの退避または復帰要求をネットワーク経由で上記第2の機器に送出するページ退避／復帰要求手段を設ける一方、上記第2の機器には、上記第1の機器のページ退避／復帰要求手段からの要求に応じて、要求されたメモリページの退避または復帰処理を上記記憶装置を用いて行うページ退避／復帰手段を設けたことを特徴とする。

【0010】このような構成において、組み込みプログラムを実行するCPUを内蔵した第1の機器、つまり組み込み機器（ノード）は、仮想記憶のメモリページの退避（または復帰）を、ネットワークを介して第2の機器（メモリサーバ）へ（または第2の機器から）行うことができるため、ハードディスク等の外部記憶装置を設けなくても、仮想記憶により大量のメモリを利用することが可能となる。

【0011】ここで、上記第1の機器内で複数のプロセスが並列動作可能な構成では、各プロセス毎にそれぞれ仮想アドレス空間（仮想メモリ）を持つことから、上記第2の機器のページ管理手段では、各メモリページを、当該ページを利用するプロセスの識別情報と共に管理するとよい。

【0012】また、上記ネットワークに接続される第1の機器が複数存在する構成では、上記各メモリページを、当該ページを利用する第1の機器の識別情報と共に管理するとよい。

【0013】また本発明は、上記ネットワークに接続される第2の機器が複数存在するシステムにおいて、上記第1の機器内に、メモリページの退避または復帰が必要な場合に、メモリページの退避先または復帰元を要求していることを、一斉通報によりネットワーク上の全ての第2の機器に通知して、その通知に対する応答に基づいて、メモリページの退避先または復帰元を決定するページ退避先／復帰元決定手段を設けたことを特徴とする。

【0014】このような構成においては、一斉通報通信を利用した要求通知により、メモリページの退避先または復帰元となる第2の機器を容易に決定することができる。ここでは、第1の機器からの一斉通報通信を利用した要求通知により決定した退避先を、その第1の機器に固有のメモリページ退避先として固定する方法、つまり退避先となる第2の機器を第1の機器毎に予め決定して

おく方法、またはメモリページの退避が必要となる度に、一斉通報通信を利用した要求通知により退避先となる第2の機器を動的に決定する方法が適用可能である。

【0015】また復帰元についても、必ず退避先の第2の機器とする方法、または上記した退避先を決定するのと同様の一斉通報通信を利用した要求通知により決定する方法が適用可能である。退避先と復帰元とを異なるようにした場合、つまりメモリページを第2の機器間で移動させた場合、第2の機器間の負荷分散を図ることができる。

【0016】この他に、第1の機器からのメモリページの退避先または第1の機器に対するメモリページの復帰元を決定するための特定の機器である第3の機器（アービトライト）を上記ネットワークに接続し、第1の機器から当該第3の機器に問い合わせをすることで、メモリページの退避先または復帰元を決定することも可能である。

【0017】

【発明の実施の形態】以下、本発明の実施の形態につき図面を参照して説明する。

【0018】〔第1の実施形態〕図1は本発明の第1の実施形態に係る仮想記憶制御システムの構成を示すブロック図である。同図において、イーサネット、ATM（Asynchronous Transfer Mode）、ファイバーチャネルなどの高速なネットワーク11には、CPU121が組み込まれた電子機器（組み込み機器）、例えば制御機器（以下、ノードと称する）12が接続されている。ノード12には、入出力インタフェース装置13を介して各種制御対象（図示せず）が接続されている。

【0019】ノード12内では、CPU121が、仮想記憶を用いた（組み込みOS及び当該OSの管理下で動作するプログラムからなる）組み込みプログラム122を実行する。ノード12には、小容量の物理メモリ（実メモリ）123が内蔵されている。またノード12には、組み込みプログラム122から見る事ができる大容量の仮想的なメモリ（仮想メモリ）124が等価的に存在する。この仮想メモリ124は、次に述べるメモリサーバ14により提供される。ノード12内で処理対象となる仮想メモリ124のメモリページのデータ（実ページデータ）は、ノード12にてCPU121により処理可能なように、実メモリ123上の任意の領域に配置される。

【0020】ネットワーク11にはまた、パーソナルコンピュータ、ワークステーション等の情報処理機器により実現されるメモリサーバ（メモリサーバ機器）14が接続されている。メモリサーバ14は、ノード12上の組み込みプログラム122に対して仮想メモリ124（上のページ）を提供するものであり、その仮想メモリ124上のページの実体が記憶される大容量の外部記憶装置140を備えている。この外部記憶装置140は、

図 1 の例のようにハードディスクを用いたディスクアレイにより実現される他、半導体ディスク、DVD-RAM などによっても実現可能である。メモリサーバ 14 は、外部記憶装置 140 に記憶（退避）されているメモリページの管理と、ノード 12 上の組み込みプログラム 122 に提供（復帰）したメモリページの管理を行うページ管理機構 141、及びページ管理機構 141 の管理のもとで、ノード 12 内の組み込みプログラム 122 から要求されたメモリページの退避または復帰処理を外部記憶装置 140 を用いて実行するページ退避／復帰部 142 を有する。

【0021】次に、図 1 の構成のシステムにおける動作を説明する。ノード 12 上の組み込みプログラム 122 の動作によって、仮想メモリ 124 上のページ（仮想記憶のメモリページ）の退避／復帰が必要になったとする。一般の、仮想記憶方式を適用するパーソナルコンピュータ等の情報処理機器であれば、ハードディスクや半導体ディスクなどの大容量記憶装置を機器の一部として備えているため、当該記憶装置を対象としてメモリページの退避／復帰を行う。ところが、本実施形態における組み込み機器としてのノード 12 は、大容量記憶装置を備えていないため、上記一般の情報処理機器のようなメモリページの退避／復帰は行えない。

【0022】そこで本実施形態では、ノード 12 内で動作する組み込みプログラム 122 は、メモリページの退避／復帰を外部のメモリサーバ 14 に要求して行う。つまり組み込みプログラム 122 はメモリページの退避／復帰の要求機能を有する。そこで以後、ノード 12 内の組み込みプログラム 122 をクライアント 122 と呼ぶ。

【0023】クライアント 122 側では、メモリページを例えば退避する場合、図 2 に示すような構造のメモリページ情報、即ち退避したいページの内容（実ページデータ）21 及び当該ページの仮想メモリ 124 上のアドレス（仮想アドレス）VA を含むページ管理情報 22 からなるメモリページ情報を、ノード 12 からネットワーク 11 を介してメモリサーバ 14 へ送信する。メモリサーバ 14 側では、受信したメモリページ情報をページ退避／復帰部 142 により外部記憶装置 140 に保存する。なお、外部記憶装置 140 にはメモリページ情報中の実ページデータ 21 のみを記憶し、ページ管理情報 22 についてはページ管理機構 141 で保持管理するようにしても構わない。

【0024】次に、メモリページを復帰する場合、クライアント 122 側では、復帰したいページの仮想アドレス VA をノード 12 からネットワーク 11 を介してメモリサーバ 14 に対して送信し、ページの内容を送るよう要求する。

【0025】メモリサーバ 14 は、クライアント 122 側から送られた仮想アドレス VA に基づき、当該クライ

アント 122 から要求されたメモリページをページ管理機構 141 により特定する。そしてメモリサーバ 14 は、要求されたメモリページの情報をページ退避／復帰部 142 により外部記憶装置 140 から取り出し、ネットワーク 11 を介して要求元のクライアント 122 へ送信する。クライアント 122 は、メモリサーバ 14 から送られた自身の要求したメモリページの情報を、ノード 12 内で CPU 121 により処理可能なように実メモリ 123 上の任意の領域に配置する。これによりノード 12 における、メモリサーバ 14 との協調動作によるメモリページの復帰処理が終了する。

【0026】ところで、クライアント 122 が複数のプロセスを並列実行する機能を有している場合、ノード 12 上では、各プロセスがそれぞれの仮想メモリ（仮想空間）124 を持つ。この場合、メモリサーバ 14 との間で授受されるメモリページの情報には、図 2 中のページ管理情報 22 に代えて、図 3 に示すようなプロセスを識別するための情報（プロセス ID）が追加されたページ管理情報 32 を持たせればよい。

【0027】〔第 2 の実施形態〕前記第 1 の実施形態は、メモリサーバとクライアントがそれぞれ 1 つずつ存在するシステムの例であるが、本発明はメモリサーバとクライアントがそれぞれ複数個存在するシステムにも応用可能である。そこで、本発明をメモリサーバとクライアントがそれぞれ複数個存在する仮想記憶制御システムに適用した第 2 の実施形態について図面を参照して説明する。

【0028】図 4 は本発明の第 2 の実施形態に係る仮想記憶制御システムの構成を示すブロック図である。同図において、（図 1 中のネットワーク 11 に相当する）ネットワーク 41 には、複数の組み込み機器、例えば 2 つの組み込み機器（ノード）42-1、42-2 と、複数のメモリサーバ、例えば 2 つのメモリサーバ 44-1、44-2 とが、それぞれ接続されている。

【0029】ノード 42-1、42-2 には、CPU 421-1、421-2、組み込みプログラム（クライアント）422-1、422-2、及び小容量の実メモリ 423-1、423-2 が組み込まれている。またノード 42-1、42-2 には、クライアント 422-1、422-2 から見ることが出来る仮想メモリ 424-1、424-2 がクライアント 422-1、422-2 で並列に実行し得るプロセス数分等価的に存在する。この仮想メモリ 424-1、424-2 は、メモリサーバ 44-1 または 44-2 により提供される。

【0030】さて、図 4 のシステムでは、クライアントが複数存在（ここでは、ノード 42-1 のクライアント 422-1 とノード 42-2 のクライアント 422-2 の 2 つが存在）するので、メモリサーバ 44-1、44-2 との間で授受するメモリページ情報には、ノード（クライアント）を識別するための情報（ノード ID）も必要となる。しかも図 4 のシステムでは、ノード 42-1、42-2

上で、各プロセスがそれぞれの仮想メモリ（仮想空間）424-1、424-2を持つことから、図3と同様にプロセスを識別するための情報（プロセスID）も必要となる。そこで本実施形態では、図3中のページ管理情報32に代えて、図5に示すようなノードIDが追加されたページ管理情報52を持たせればよい。

【0031】また、図4のシステムでは、メモリサーバが複数存在（ここでは、メモリサーバ44-1及び44-2の2つが存在）するため、当該システムには、あるクライアントからのメモリページ退避先をどのサーバにするかを決定する機構が必要となる。このメモリページ退避先の決定には、退避先となるメモリサーバを、クライアント毎にそれぞれ1つ予め定めておく第1の退避先決定方法と、退避先となるメモリサーバを退避の度に動的に決定する第2の退避先決定方法のいずれかが適用可能である。

【0032】ここで、上記第1の退避先決定方法の手順について、ノード42-1のクライアント422-1からのメモリページ退避先を決定する場合を例に、図6のシーケンスチャートを参照して説明する。

【0033】まずクライアント422-1は、メモリページの退避先となるメモリサーバを要求していることを、ネットワーク41上の全てのメモリサーバに同時に通知する（ステップS1）。この通知は、ネットワーク41が例えイーサネットではIP（Internet Protocol）と呼ばれるプロトコルを適用しているならば、broadcast（一斉同報）という仕組みを使えば、容易に実現できる。

【0034】各メモリサーバ44-1、44-2は、それぞれの状況に応じた応答（退避先サーバ通知または非退避先サーバ通知）をクライアント422-1に返す。ここでは、メモリサーバ44-2からは自身がクライアント422-2からのメモリページの退避先のサーバとなることを通知する退避先サーバ通知が返されて（ステップS2）、メモリサーバ44-1からは自身がクライアント422-2からのメモリページの退避先のサーバとならないことを通知する非退避先サーバ通知が返されるものとする（ステップS3）。

【0035】これを受けてクライアント422-1は、メモリページの退避先のメモリサーバを選ぶ。ここでは、メモリサーバ44-2が退避先サーバとして選ばれる。以後、クライアント422-1は、メモリページを退避する必要が発生する都度、図5に示すような構造のメモリページ情報を、ノード42-1からメモリサーバ44-2へ送信して、対応するメモリページの退避を要求する。

【0036】以上の説明では、メモリページの退避先となるメモリサーバは勿論、退避先とならないメモリサーバも応答を返すものとしているが、退避先とならないメモリサーバからは応答を返さない構成であっても構わない。

【0037】次に上記第2の退避先決定方法について説明する。この第2の退避先決定方法では、上記第1の退避先決定方法で適用した、クライアント422-1から各サーバへの一斉同報による退避先メモリサーバ要求を、予めではなくて、メモリページの退避が必要となる都度送出する。したがって、第2の退避先決定方法では、メモリページの退避の度に退避先のメモリサーバが決定されることになる。

【0038】なお、図4のシステムに、図7に示すように、メモリページの退避元（または復帰元）となるクライアントと退避先（または復帰元）となるメモリサーバの組を決定する機能を有する特別なサーバであるアービトライト70を追加して、各クライアント422-1、422-2から当該アービトライト70に対してメモリページの退避先（または復帰元）となるメモリサーバを問い合わせる構成とすることも可能である。

【0039】この図6のシステムでも、前記第1の退避先決定方法と同様の手法、即ち予めクライアントからアービトライト70に要求して退避先メモリサーバを決定しておく手法と、前記第2の退避先決定方法と同様の手法、即ちメモリページの退避が必要となる都度クライアントからアービトライト70に要求して退避先メモリサーバを動的に決定する手法のいずれも適用可能である。

【0040】次に、図4または図6のシステムにおけるメモリページの復帰について、クライアント422-2がメモリページを復帰したい場合を例に説明する。

【0041】まず、図4または図6のシステムでは、メモリサーバが複数存在するため、クライアント422-1がメモリページを復帰したい場合にも、先の退避先のメモリサーバを決定するのと同様に、どのサーバから復帰すればよいのか、つまり復帰元メモリサーバを決定する必要がある。このメモリページの復帰元メモリサーバの決定には、必ず退避先サーバから復帰すると定めておく第1の復帰元決定方法と、メモリページの復帰が必要となる都度、復帰元となるメモリサーバを動的に決定する第2の復帰元決定方法のいずれかが適用可能である。

【0042】上記第1の復帰元決定方法は最も処理が単純となるもので、クライアント422-1は単に退避先となったメモリサーバを記憶しておくだけでよい。但し、この手法を適用した場合、システムの状況によっては、特定のメモリサーバに大きな負荷がかかるという不具合がある。

【0043】上記第2の復帰元決定方法は、この第1の復帰元決定方法の不具合を解決するものであり、メモリサーバ間で負荷分散を図るため、その負荷状況等に応じてメモリページがメモリサーバ間で移動すること、つまり退避先メモリサーバと復帰元メモリサーバとが必ずしも一致しないことを前提としている。

【0044】この第2の復帰元決定方法を図4のシステムに適用した場合には、前記第2の退避先決定方法と同

様の手法により、メモリページの復帰が必要となる都度、ネットワーク41上の全てのメモリサーバ(44-1, 44-2)に一斉同報により復帰元メモリサーバ要求を送出して、その応答から復帰元メモリサーバを決定すればよい。

【0045】また、図6のシステムに上記第2の復帰元決定方法を適用した場合には、メモリページの復帰が必要となる都度、アービトレイタ70に復帰元メモリサーバを問い合わせ、その応答から復帰元メモリサーバを決定すればよい。

【0046】

【発明の効果】以上詳述したように本発明によれば、組み込み機器において、ハードディスク等の外部記憶装置を設けることなく、仮想記憶により大量のメモリを利用できるようになる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係る仮想記憶制御システムの構成を示すブロック図。

【図2】同実施形態で適用されるメモリページ情報のデータ構造例を示す図。

【図3】クライアント(組み込みプログラム)が複数のプロセスを並列実行する場合のメモリページ情報のデータ構造例を示す図。

【図4】本発明の第2の実施形態に係る仮想記憶制御シ

ステムの構成を示すブロック図。

【図5】同実施形態で適用されるメモリページ情報のデータ構造例を示す図。

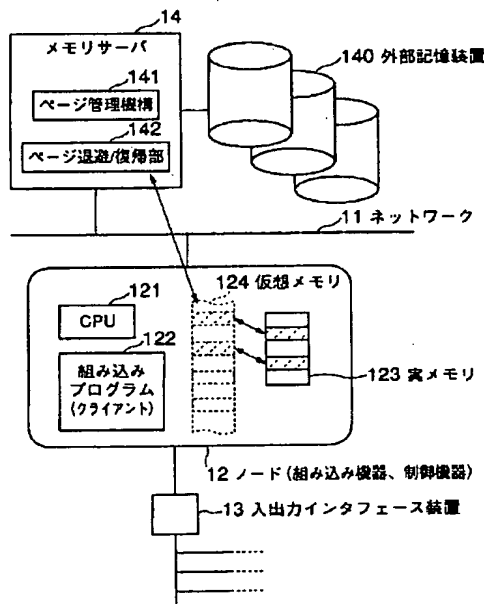
【図6】同実施形態で適用される第1の退避先決定方法の手順について、ノード42-1のクライアント422-1からのメモリページ退避先を決定する場合を例に説明するためのシーケンスチャート。

【図7】図4の構成の変形例を示すブロック図。

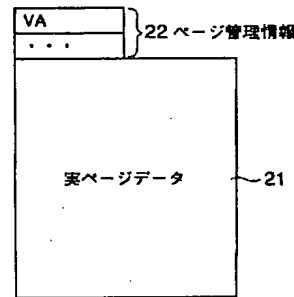
【符号の説明】

- 11, 41…ネットワーク  
 12, 42-1, 42-2…ノード(組み込み機器、第1の機器)  
 14, 44-1, 44-2…メモリサーバ(第2の機器)  
 22, 32, 52…ページ管理情報  
 70…アービトレイタ(第3の機器)  
 121, 421-1, 421-2…CPU  
 122, 422-1, 422-2…組み込みプログラム(クライアント、ページ退避/復帰要求手段、ページ退避先/復帰元決定手段)  
 20 123, 423-1, 423-2…実メモリ  
 124, 424-1, 424-2…仮想メモリ  
 141…ページ管理機構  
 142…ページ退避/復帰部

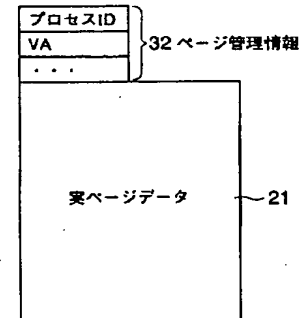
【図1】



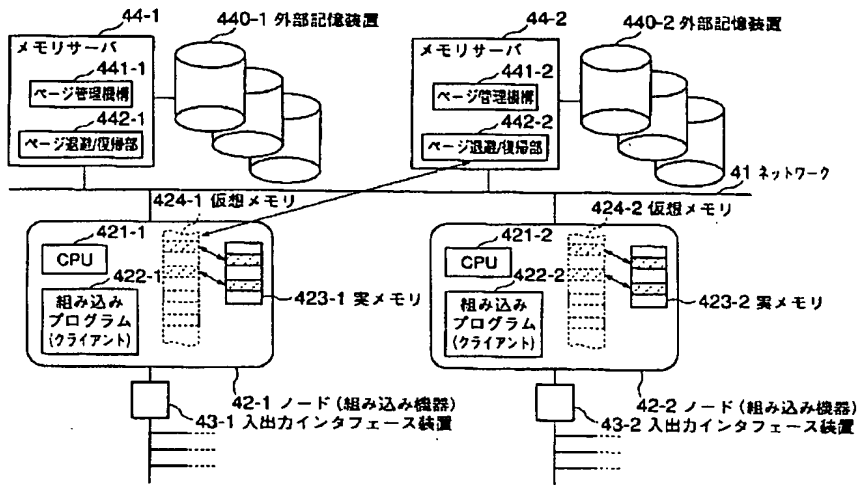
【図2】



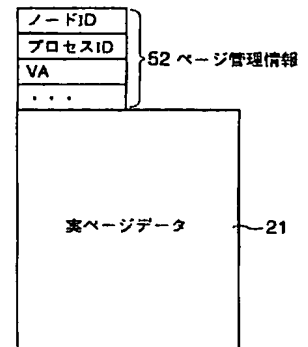
【図3】



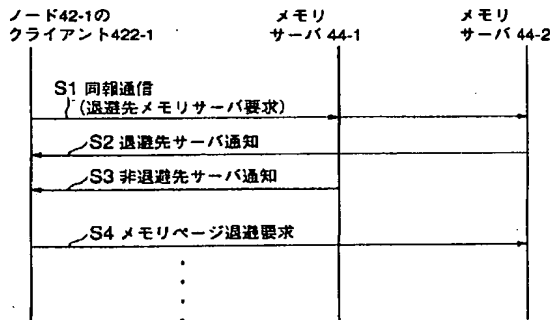
【図4】



【図5】



【図6】



【図7】

